



Herkko Salonen

ÄÄNEEN LUKEMISEN SUJUVUUDEN AUTO- MAATTINEN ARVIOINTI

Kandidaatintyö
Informaatioteknologian ja viestinnän tiedekunta
Tarkastaja: Okko Räsänen
Joulukuu 2019

TIIVISTELMÄ

Herkko Salonen: Ääneen lukemisen sujuvuuden automaattinen arviointi
Kandidaatin työ
Tampereen yliopisto
Tietotekniikan tutkinto-ohjelma
Joulukuu 2019

Lukivaikeuksista kärsiviä etsitään arvioimalla ääneen lukemistaan. Arviointi vie aikaa, joka voitaisiin käyttää lukivaikeuden hoitamiseen. Automaattinen arviointi lisäisi hoitoon käytettävissä olevia resursseja, ja sitä voitaisiin käyttää myös harjoittelun tukena. Tässä työssä esitellään järjestelmä, joka arvioi lasten ääneen lukemisen sujuvuutta. Järjestelmä arvio puheen sujuvuutta joko yleisesti tai painotuksen, sujuvuuden, tahdin ja tunteilmaisun suhteen. Työssä analysoidaan myös järjestelmässä käytettyjen piirteiden vaikutusta arviointikriteereihin.

Avainsanat: lukutaidon arviointi, lukusujuvuus, prosodia, automaattinen arviointi

Tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck –ohjelmalla.

ABSTRACT

Herkko Salonen: Automatic evaluation of read speech fluency
Bachelor's Thesis
Tampere University
Bachelor's Degree Program in Computer Science
December 2019

Dyslexia patients are screened by evaluating people's oral reading. Evaluation takes time which, could be used for treating dyslexic patients. Automatic evaluation of reading could add resources for treating patients, and it could also be used as a reading companion tool. In this paper an automatic evaluation system for children's oral reading fluency is introduced. The system evaluates reading fluency either overall or regards to phrasing, smoothness, pace, and expression. Correlations between used speech features and fluency criteria are also discussed.

Keywords: reading evaluation, reading fluency, prosody, automatic evaluation

The originality of this thesis has been checked using the Turnitin OriginalityCheck service.

ALKUSANAT

Haluan kiittää ohjaajaani Okko Räsästä hyvistä vinkeistä ja saamastani tuesta työn aikana. Kiitän lisäksi Joni Kämäräistä, jonka pitämät luennot ovat motivoineet minua näiden vuosien aikana. Kiitokset kuuluvat myös Jarkko Hautalalle ja hänen tutkimusryhmälleen aineiston ja arvioiden luovuttamisesta työtä varten.

Tampereella, 12.12.2019

Herkko Salonen

SISÄLLYSLUETTELO

1. JOHDANTO	1
2. LUKEMISEN SUJUVUUDEN ARVIOINNIN TAUSTA.....	3
2.1 Sujuvuuden perinteiset arviointikriteerit	3
2.2 Sujuvuuden automaattinen arviointi.....	4
3. MENETELMÄT.....	5
3.1 Työssä käytetty data	5
3.2 Datan esikäsittely	6
3.2.1 Automaattinen puheentunnistus	6
3.2.2 Perussävelen tunnistus	6
3.3 Piirteiden laskeminen	7
3.4 Tilastolliset menetelmät.....	9
3.4.1 Lineaarinen regressiomalli	9
3.4.2 Muut tilastolliset mallit	10
3.5 Piirteiden valitseminen	10
3.6 Automaattisen järjestelmän arviointi	10
4. TULOKSET	12
5. YHTEENVETO.....	20
LÄHTEET	21

LIITE A: AINEISTOSSA KÄYTETYT TARINAT

LIITE B: AINEISTON ARVIOINTIIN KÄYTETYT KRITERIT

LYHENTEET JA MERKINNÄT

ASR	engl. Automatic Speech Recognition, automaattinen puheentunnistus
CORFS	engl. Comprehensive Oral Reading Fluency Scale, kattava ääneen lukemisen sujuvuuden asteikko
F0	perussävel
MAE	engl. Mean Absolute Error, absoluuttinen poikkeama
MFS	engl. Multidimensional Fluency Scale, moniulotteinen sujuvuuden asteikko
MLP	engl. Multilayer Perceptron, eteenpäin kytketty monikerroksinen perseptroni-verkko. Yksinkertainen neuroverkkorakenne.
NAEP	engl. National Assessment of Educational Progress, kansallinen arvio koulutuksen etenemisestä
RMSE	engl. Root Mean Squared Error, keskineliövirheen neliöjuuri
R^2	selitysaste, engl. coefficient of determination
SVM	engl. Support-Vector Machine, tukivektori-kone
WCPM	engl. Words Correct Per Minute, sanoja oikein minuutissa
WPM	engl. Words Per Minute, sanoja minuutissa

1. JOHDANTO

Lukutaito on nykyajan yhteiskunnassa elämisen edellytys. Lukutaidottomuus heikentää mahdollisuuksia jatkokoulutukseen ja lisää syrjäytymisriskiä [28]. Pelkkä lukutaito ei kuitenkaan riitä menestymiseen; pitää myös ymmärtää lukemansa. Useat tutkimukset osoittavat, että lukemisen sujuvuus korreloi luetun ymmärtämisen kanssa [4][36]. Suomi on maailmanlaajuisesti lukutaidon kärkimaita, mutta silti 11 % suomalaisista aikuisista on lukutaidoltaan heikkoja [26]. Yleisen lukutaidon kehittämiseksi on tärkeä löytää lukivaikeuksista kärsivät aikaisin.

Lukemisen sujuvuutta tyypillisesti mitataan sanastollisten ja prosodisten piirteiden avulla. Sanastollisiin piirteisiin kuuluvat muun muassa lukunopeus ja lukemisen tarkkuus, eli kuinka moni luettu sana on luettu oikein. Prosodisiin piirteisiin kuuluvat esimerkiksi tavujen painotus, intonaatio ja tauotus. Tutkimuksien mukaan prosodiset piirteet kuvaavat paremmin luetun ymmärtämistä. Etenkin haastavissa teksteissä prosodiset piirteet soveltuvat myös luetun ymmärtämisen mittaamiseen. [4][24]

Alakouluikäisten lukemisen kehitystä seurataan tarkkaan. Tukea tarvitsevia lapsia seulotaan lukutesteillä, joissa lapset käyvät lukemassa erityisopettajalle. Erityisopettaja arvioi lukevatko lapset riittävän sujuvasti ikäänsä nähden. Sujuvuuden automaattinen arviointi voisi vähentää erityisopettajien seulonnoista aiheutuvaa työmäärää, jolloin aikaa jäisi enemmän lukemisharjoitusten pitämiseen. Automaattista sujuvuuden arviointia voidaan hyödyntää myös lukemisen harjoittelua tukevissa peleissä. Tällä hetkellä automaattista arviointia käytetään muun muassa LukiMat-hankkeen osana kehittyssä Ekapeli-Sujuvuus -pelissä. Pelissä sujuvuuden arviointi perustuu kyllä/ei-kysymyksiin, joihin vastataan hiljaisesti luettujen väittämien perusteella. [38]

Suomenkielelle kehitettyjä lukupuhunnan sujuvuuden arviointimenetelmiä ovat LukiMatin Luksun lisäksi Diagnostiset testit 1, YTTE ja Armi 1 ja 2, FinRa ry:n lukemisen testit ja Lukilasse. Diagnostiset testit 1 mittaa sujuvuutta lyhyiden ja pitkien sanojen lukemiseen käytetyn ajan perusteella [30]. YTTE-menetelmässä [15], Armi 1 [19] ja 2-testeissä [20] ja FinRa ry:n lukutesteissä [13] sujuvuutta mitataan tarinan lukunopeuden ja virheiden perusteella. Lukilasse-testissä [14] sujuvuutta mitataan oikein luettujen sanojen määrällä rajallisessa ajassa. Kaikki testit ovat manuaalisia ja FinRa ry:n testiä lukuun ottamatta normitettuja, mikä mahdollistaa tulosten vertailun valtiontasolla. Yksikään testeistä ei mittaa puheen prosodisia ominaisuuksia, vaan testit keskittyvät lukunopeuteen ja tarkkuuteen.

Tässä työssä esitellään järjestelmä, joka arvioi automaattisesti ääneen lukemisen sujuvuutta puhe-signaalin piirteistykseen avulla. Sujuvuutta arvioidaan lähinnä prosodisten piirteiden avulla. Työssä on hyödynnetty Aalto-ASR-puheentunnistinta [1] ja PiENet-äänien korkeuden arviointityökalua [29].

Näiden ohjelmien tuottamien syötteiden avulla lasketaan akustisesta puhesignaalista piirteitä, joiden perusteella pyritään arvioimaan ääneen lukemisen sujuvuutta. Sujuvuuden arviointi tapahtuu koneoppimisalgoritmin avulla, jonka opettamiseen käytetään käsin arvioituja äänitteitä. Äänitteet on arvioitu käyttäen apuna neljää eri arviointikriteeriä. Työssä käydään läpi eri puhepiirteiden korrelaatiota näihin arviointikriteereihin.

Luvussa 2 tutustutaan aiemmin tehtyihin tutkimuksiin aiheesta, esitellään niissä käytettyjä menetelmiä ja arvioidaan tutkimusten tuloksia. Luvussa 3 esitellään tässä työssä käytetyt menetelmät. Luvussa kerrotaan, miten dataa esikäsiteltiin, mitä piirteitä datasta laskettiin ja kuinka piirteitä hyödynnettiin koneoppimisalgoritmissa. Luvussa 4 esitellään tutkimustulokset ja analyysi niistä. Lopuksi luvussa 5 on yhteenveto tästä työstä.

2. LUKEMISEN SUJUVUUDEN ARVIOINNIN TAUSTA

Automaattinen ääneen lukemisen sujuvuuden mittausjärjestelmä on tullut ajankohtaiseksi 2010-luvulla tietokoneiden yleistymisen, ja prosessointikyvyn kasvamisen myötä. Järjestelmiä on kehitetty etupäässä alakouluikäisten lasten lukemisen harjoittelun tueksi [22][23][38] ja dysleksiasta kärsivien tunnistamiseksi [6][7]. Nämä järjestelmät yleensä arvioivat lasten ääneen lukemista joko kestoaltaan tai sanamäärältään rajatun tarinan perusteella. Näiden järjestelmien lisäksi automaattista puheen arviointia on käytetty vieraankielen osaamisen arvioimiseen [5].

Luvussa 2.1 esitellään perinteisiä ääneen lukemisen sujuvuuden arviointikriteerejä. Luvussa 2.2 käydään läpi aikaisemmissa julkaisuissa käytettyjä automaattisia arviointimenetelmiä.

2.1 Sujuvuuden perinteiset arviointikriteerit

Lukemisen sujuvuuden arvioinnissa ylivoimaisesti suosituin kriteeri on ollut oikein luettujen sanojen määrä minuutissa (Words Correct Per Minute, WCPM) sellaisenaan tai sovellettuna [6][7][12][22]. Useiden tutkijoiden mukaan lukemisen sujuvuuden kriteerejä ovat nopeus, tarkkuus ja ilmaisu [32][34]. WCPM mittaa näistä kahta, ja sitä on käytetty usein ainoana ääneen lukemisen sujuvuuden mittarina. Joidenkin tutkijoiden mukaan WCPM ei kuitenkaan anna riittävän laajaa kuvausta sujuvuudesta [3].

Yhdysvalloissa kehitetty NAEP-asteikko (National Assessment of Educational Progress, kansallinen arvio koulutuksen etenemisestä) on nelitasoinen ja se ottaa huomioon nopeuden lisäksi lukijan ilmaisun [27]. Asteikon tasot erottavat toisistaan lukijan yhdessä lukemien sanojen määrä. Ensimmäisellä tasolla lukija lukee sana kerrallaan, toisella tasolla lukija lukee pääasiassa kahden sanan ryhmissä, kolmannella tasolla lukija lukee useamman sanan ryhmissä, ja neljännellä tasolla luetut sanat muodostavat merkityksellisen kokonaisuuden. NAEP-asteikon automaattinen arviointi ei ole yhtä suoraviivainen menetelmä kuin WCPM:n.

Moniulotteinen sujuvuusasteikko (Multidimensional Fluency Scale, MFS) arvioi lukemisen ominaisuuksia erillään toisistaan. Asteikolla arvioidaan sanojen painotusta, sujuvuutta (tahdin tasaisuus) ja tahtia. Jokainen ominaisuus arvioidaan numeerisella 1–4 asteikolla. [41] NAEP-asteikon tavoin MFS-asteikon automaattinen arviointi vaatii usean piirteen käyttämistä, koska asteikko ei koostu vain yhdestä mitattavasta suureesta.

Benjamin et al. kehittämä CORFS-asteikko (Comprehensive Oral Reading Fluency Scale, kattava ääneen lukemisen sujuvuuden asteikko) perustuu spektrografisiin piirteisiin kuten äänenkorkeuteen ja taukojen määrään ja pituuteen. Asteikko korostaa prosodisien piirteiden merkitystä enemmän kuin muut arviointiasteikot. [3]

2.2 Sujuvuuden automaattinen arviointi

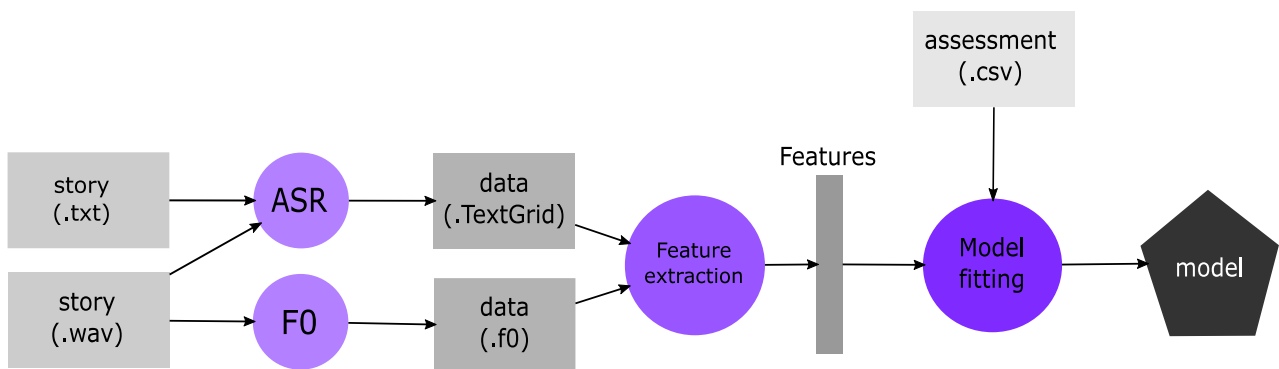
WCPM:n mittaaminen automaattisesti on mahdollista toteuttaa pelkän ASR:n (Automatic Speech Recognition, automaattinen puheentunnistin) avulla [6][7]. Tällöin arvioinnin tarkkuus riippuu automaattisen puheentunnistimen tarkkuudesta. Lasten puheen tunnistaminen on vaikeampaa kuin aikuisten puheen [16][21]. Automaattisen puheentunnistimen toiminnan edellytys on, että tunnistimen opettamisessa on käytetty lasten puheesta koostuvaa korpusta [6][7].

Bolaños et al. artikkelissa [6] kuvataan FLORA-järjestelmä, joka automaattisesti arvioi ääneen lukemisen sujuvuutta pohjautuen neliasteiseen NAEP-asteikkoon. Järjestelmässä sujuvuuden arviointi perustuu NAEP-asteikon perusteella toteutettujen piirteiden laskemiseen puhesignaalista. Sujuvuuden arviointi tapahtuu kolmen puhesignaalin piirteillä opetetun SVM-luokittelijan avulla. Järjestelmä antaa saman arvion kuin ihminen yli 76 prosentissa tapauksista. Pelkästään WCPM-kriteerin perusteella järjestelmä antaa saman arvion kuin ihminen 70 prosentissa tapauksista. [6]

Sujuvuuden arviointiin on käytetty myös mallipohjaisia ratkaisuja. Sabun ja Raon julkaisussa [35] esitellään järjestelmä, jossa prosodian arviointia varten on opetettu kielimalleja. Jokaiselle arvioinnissa käytetylle tarinalle on opetettu oma kielimallinsa, jossa sanan esiintymistodennäköisyyden määrittämiseen käytetään kahta edeltävää sanaa. Kielimallin sanoille on opetusdatan perusteella määritetty prosodisia piirteitä, kuten perussävel, pituus ja intensiteetti ajanfunktioina. Järjestelmä arvioi lukemisen prosodian ominaisuuksia vertaamalla opetettuihin kielimalleihin. [35]

3. MENETELMÄT

Tässä luvussa esitellään työssä käytetyt menetelmät ääneen lukemisen automaattiseen arviointiin. Työssä käytetty datan käsittelyn kanava mukailee aikaisemmissa tutkimuksissa käytettyjä prosesseja. Luvussa 3.1 kerrotaan työssä käytetystä aineistosta. Luvussa 3.2 käsitellään äänitiedostojen esikäsittelyä ja muuttamista helpommin käsiteltävään muotoon. Luvussa 3.3 esitellään piirteiden laskemisen menetelmät. Luvussa 3.4 kerrotaan työssä käytetyistä tilastollisista malleista. Luvussa 3.5 esitellään piirteiden arviointimenetelmät. Lopuksi luvussa 3.6 kerrotaan mallien arviointimenetelmistä.



Kuva 1: Lohkokaavio mallin opetuksesta

Kuvassa 1 on esitelty arviointijärjestelmän opettaminen lohkokaaviona. Aluksi puheentunnistimella (ASR) etsitään äänitiedostoista lauseiden, sanojen ja äänteiden aikaleimat, ja perussävelentunnistimella selvitetään puhujan äänenkorkeus (F0) äänitteen aikana. Tunnistimien tuloksista lasketaan joukko lukemisen sujuvuutta kuvaavia piirteitä (engl. *feature extraction*). Lasketut piirteet sovitetaan ihmisen tekemien arvioiden kanssa tilastolliseksi malliksi (engl. *model fitting*). Järjestelmä mukailee soveltaen Bolaños et al. julkaisussa [6] esiteltyä FLORA-järjestelmää.

3.1 Työssä käytetty data

Työssä käytettiin Niilo Mäki Instituutin ReadDrama-tutkimuksessa [33] kerättyä aineistoa. Aineisto koostuu 106 eri lapsen lukemista 210 äänitteestä. Jokainen äänite sisältää yhden 150-sanaisen tarinan kolmesta vaihtoehtoisesta tarinasta. Tarinat ovat saatavilla työn liitteessä A. Työssä äänitteistä käytettiin 207 kappaletta, koska kaikkia äänitteitä ei oltu käsin arvioitu sujuvuuden osalta.

Äänitteiden lukijat olivat 3.–4.-luokkalaisia oppilaita. Oppilaat lukivat tarinat ensin äänettömästi, jonka jälkeen he vastasivat monivalintatehtävään, joka mittasi luetun ymmärtämistä. Tämän jälkeen oppilaat lukivat tarinan uudelleen ääneen. Äänetön lukeminen ja monivalintatehtävän tarkoitus oli parantaa oppilaan ilmaisukykyä ääneen lukemisvaiheessa. [33]

Äänitteiden arvioinnissa käytetyt kriteerit pohjautuvat moniulotteiseen sujuvuusasteikkoon [40], joka esiteltiin luvussa 2.1. Asteikkoa on sovellettu laajentamalla pisteytystä viisiportaiseksi. Sen lisäksi asteikkoon lisättiin tunneilmaisua kuvaava ulottuvuus. Tarkemmat arviointikriteerit ovat saatavilla liitteessä B.

3.2 Datan esikäsittely

Alkuperäisestä datasta työssä käytettiin 207 pakkaamatonta wav-tiedostoa. Äänitteet oli näytteistetty 44,1 kHz taajuudella ja 16 bitin tarkkuudella. Käsittelyä varten äänitteet alinäytteistiin 16 kHz näyteistystaajuudelle.

Työssä käytettiin kohinasuodatettuja äänitteitä, jotta automaattinen puheentunnistus ja perussävelen tunnistaminen onnistuisivat paremmin. Kohinasuodatus oli tehty äänitteille keräyksen yhteydessä.

3.2.1 Automaattinen puheentunnistus

Piirteiden laskeminen edellyttää sanojen ja äänteiden tunnistamista äänitteistä. Työssä käytettiin AaltoASR-puheentunnistinta [17] virkkeiden, sanojen ja äänteiden tunnistamiseen äänitteistä. Tähän käytettiin puheentunnistimen forced alignment -ominaisuutta. Ominaisuuden käyttämisen tarkoitus oli parantaa puheentunnistimen tuloksia lasten puhetta tunnistettaessa. Ominaisuus etsii äänitteestä annetun tekstitiedoston sanoille aikaleimat. Sanat muutetaan suomen kielen ääntämismallin avulla todennäköisimmäksi äännesekvenssiksi. Tämän sekvenssin kautta kohdistusalgoritmi valitsee aikaleimat erikseen ääniteille.

Kohdistus algoritmi perustuu Viterbin algoritmiin [40], jonka on tarkoitus löytää todennäköisin sekvenssi Markovin piilomalleja. Tässä yhteydessä Viterbin algoritmia käytetään todennäköisimmän äännesekvenssin löytämiseen suomen kielen ääntämismallin avulla.

Työssä käytetty forced alignment -kohdistus ei toimi kunnolla epäsujuvan puheen tunnistamisessa. Saatavilla ei kuitenkaan ollut parempia ASR-malleja, jotka olisi tarkoitettu lasten puheen tunnistamiseen ja/tai epäsujuvan puheen tunnistamiseen. Työssä haluttiin kokeilla, kuinka hyvin saatavilla olevilla työkaluilla sujuvuusongelmaa pystytään ratkomaan.

3.2.2 Perussävelen tunnistus

Puheentunnistamisen kautta saatujen tietojen ohella puheen perussävelen taajuus mahdollistaa puhe-signaalia kuvaavien piirteiden laskemisen. Tonaalisissa kielissä kuten mandariinikiinassa perus-

sävel vaikuttaa sanan merkitykseen. Suomenkielessä perussävelen muutokset muuttavat vain sanoman sisällön painotusta. Perussävelen arvoilla on sitä kautta yhteys ääneen lukemisen tunneilmaisuuun. Työssä käytetyn perussävelen tunnistusohjelma on esitelty Airaksinen et al. artikkelissa [2][29].

Perussävelen tunnistus tuottaa jokaista äänitettä vastaavan f0-tiedoston, jossa sisältää vektorin perussävelen arvoja 10 ms aikaresoluutiolla. Tunnistin tuottaa perussävelen arvon vain soinnillisille äänteille, koska soinnittomilla äänteillä kuten /p/ ja /t/ ei ole perustajuutta [18]. Tunnistus tuottaa kuitenkin tarpeeksi arvoja jokaista sanaa kohden, jotta dataa voidaan käyttää piirteiden laskemisessa.

3.3 Piirteiden laskeminen

Työssä käytetyistä piirteistä muutamat ovat samoja, mitä Bolaños et al. artikkelissa [6] esiteltiin. Artikkelissa esiteltiin järjestelmä, joka mittasi lasten ääneen lukemisen sujuvuutta piirteisiin pohjautuen. Artikkelissa esiteltyt piirteet perustuva NAEP-asteikkoon [27]. Kaikkia artikkelissa esiteltyjä piirteitä ei toteutettu tässä työssä niiden toteuttamisen vaikeuden vuoksi. Forced alignment -kohdistus ei tuottanut niin monipuolista dataa kuin artikkelissa esitellyssä järjestelmässä käytetty puheentunnistin. Kohdistuksen tuottaman datan perusteella on esimerkiksi hyvin vaikeaa määritellä, onko sana lausuttu oikein. Taulukossa 1 on esitelty kaikki työssä käytetyt piirteet.

Taulukko 1: Työssä käytetyt piirteet

PIIRTEEN NIMI	PIIRTEEN KUVAUS
TAUK	Äänteiden välissä olevien taukojen määrä.
TAUK _{KA}	Äänteiden välissä olevien taukojen keskipituus.
TAUK _{MAX}	Pisimmän äänteiden välissä olevan tauon pituus.
TAUK _{KH}	Äänteiden välissä olevien taukojen pituuden keskihajonta.
VOK _{KA}	Vokaaliäänteiden keskipituus. Peräkkäisten vokaalien pituudet laskettiin yhteen.
VOK _{MAX}	Pisimmän vokaaliäänteen pituus. Peräkkäisten vokaalien pituudet laskettiin yhteen.
VOK _{KH}	Vokaaliäänteiden pituuden keskihajonta. Peräkkäisten vokaalien pituudet laskettiin yhteen.
KONS _{KA}	Konsonanttiäänteiden keskipituus.
KONS _{MAX}	Pisimmän konsonanttiääneen pituus.
KONS _{KH}	Konsonanttiäänteiden pituuden keskihajonta.
VIRK _{KA}	Virkkeiden välisten taukojen keskipituus.
LOPPU	Puheentunnistimen tunnistaman viimeisen hiljaisuuden pituus. Arvio puheentunnistimen tekemien virheiden määrälle.
F0 _{KA}	Piirre kuvaa, kuinka paljon perussävel keskimäärin vaihtelee sanojen sisällä.

$F0_{KH}$	Piirre kuvaa, kuinka paljon sanojen keskimääräinen perussävel vaihtelee.
WPM	Piirre on luettujen sanojen määrä minuutissa. Piirre ei huomio, onko sanat luettu oikein.
WPM_{VAR}	Piirre kuvaa, kuinka paljon virkkeiden välillä on lukunopeuden vaihtelua.

TAUK-piirteet käsittelevät kaikkia äänteiden välisiä taukoja, jotka puheentunnistin on tunnistanut. Nämä piirteet lukuun ottamatta taukojen pituuden keskihajontaa esiintyvät myös Bolaños et al. artikkelissa. Bolañosin mukaan Clayn ja Imlachin [8] artikkelissa todetaan heikkojen lukijoiden pitävän enemmän taukoja ja heidän pitämät taukonsa ovat pidempiä kuin taitavien lukijoiden pitämät. Arviointikriteerien mukaan taitava lukija lukee tasaiseen tahtiin, ja heikommilla lukijoilla esiintyy lukunopeuden vaihtelua. Taukojen pituuden keskihajonta kuvaa arviointikriteerien painotusta, sujuvuutta, sekä jossain määrin tahtia.

Työssä käytettyjen arviointikriteerien mukaan heikot lukijat liu’uttavat, takeltelevat tai toistavat luke- maansa. Tämä viittaa heikkoon tekstin dekodaus taitoon [39]. Lukija toistaa lukemaansa, koska ei kykene etenemään tekstissä. Toisto voi esiintyä soinnillisten äänteiden venyttämisenä. Soinnillisten äänteiden pituuden mittaamiseksi tuotettiin VOK- ja KONS-piirteet. Nämä piirteet mittaavat vokaalien ja konsonanttien keskipituutta, keskipituuden hajontaa ja pisimpien vokaali- ja konsonanttiaänteiden pituutta. Piirteiden on tarkoitus mitata eritoten arviointikriteerien painotusta ja sujuvuutta.

Rasinskin mukaan etevä lukija kiinnittää huomiota välimerkkeihin, etenkin virkkeiden välillä [31]. Lukiessa pidettävien hengähdystaukojen tulisi esiintyä välimerkkien kohdalla. Heikommat lukijat joutuvat pitämään useammin hengähdystaukoja, ja he kiinnittävät vähemmän huomiota välimerkkeihin. Piirteen $VIRK_{KA}$ on tarkoitus mitata lukemisen painotusta ja tahtia.

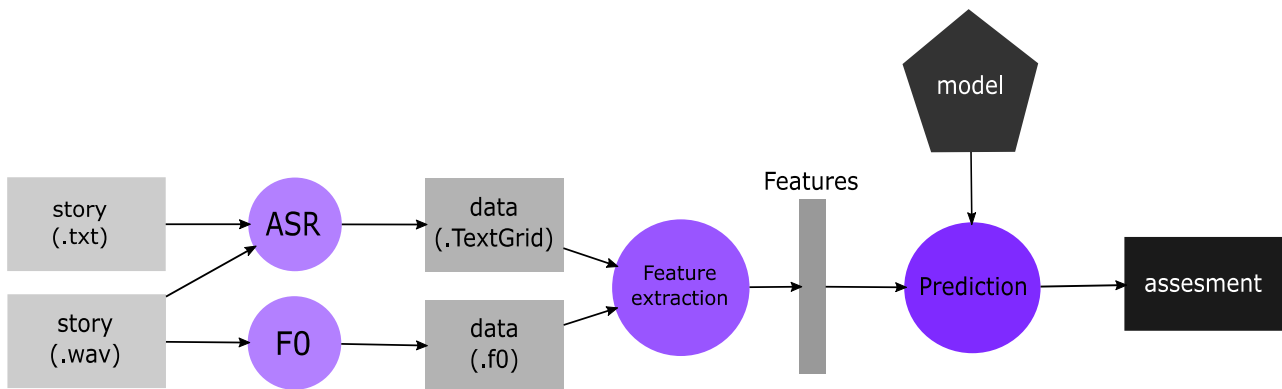
Puheentunnistuksesta saatujen aikaleimojen avulla jokaista sanaa vastaavat perussävelen arvot toimivat pohjana piirteille $F0_{KA}$ ja $F0_{KH}$. Työssä käytettyjen arviointikriteerien mukaan hyvää tunneilmaisua kuvaa eläytyminen tekstiin. Tämä ilmenee äänensävyyn muuttumisena tekstin sisällön mukaan. Piirteiden $F0_{KA}$ ja $F0_{KH}$ tarkoitus on mitata tunneilmaisua.

Piirre WPM (engl. *words per minute*, sanoja minuutissa) kuvaa lukemisen nopeutta. WPM on miltei sama suure kuin WCPM, jota Bolaños et al. [6] mukaan pidetään yleisesti todella kuvaavana piirteenä. Näiden suureiden ero on se, ettei WPM mittaa lukemisen tarkkuutta. Luvussa 2 todettiin, että lukemisen sujuvuus muodostuu nopeudesta, tarkkuudesta ja ilmaisusta, ja WPM mittaa yhtä näistä. Lisäksi työssä käytettyjen arviointikriteerien mukaan lukunopeus vaikuttaa arviointiin painotuksen, sujuvuuden ja tahdin osalta. Täten lukunopeus on merkittävä ääneen lukemisen sujuvuuden arviointin kriteeri. Lukunopeuden vaihtelu, jota piirre WPM_{VAR} mittaa, on ominaista keskitason lukijoille arviointikriteerien mukaan. Arviointikriteerien mukaan keskitasoinen lukija lukee helpot lauseet nopeasti, mutta hidastaa tahtiaan vaikeissa lauseissa.

3.4 Tilastolliset menetelmät

Äänitteistä laskettujen piirteiden sekä aineiston mukana toimitettujen sujuvuusarvioiden perusteella aineistoon sovitettiin useita tilastollisia malleja. Aluksi selvitettiin kuinka tarkkaan sujuvuuden arviointiin päästään käyttämällä kaikkia piirteitä ja lineaarista regressiomallia. Tämän jälkeen eri piirrekombinaatioiden tuottamaa tarkkuutta vertailtiin lineaarisen regression avulla. Lopuksi vertailtiin arvioinnin tarkkuutta muilla suosituilla tilastollisilla malleilla.

Tilastollisien mallien tuottamiseen käytettiin apuna Pythonin scikit-learn -kirjastoa. Scikit-learn on helppokäyttöinen ja suosittu koneoppimiseen tarkoitettu kirjasto. Kirjastoa käytetään yli 40 % kaikista GitHubin koneoppimisprojekteista [10]. Tilastollisien mallien toteutuksien yksityiskohdat ovat saatavissa kirjaston kotisivulla. [37]



Kuva 2: Lohkomalli tilastollisen mallin testaamisesta

Kuvassa 2 esitelty lohkomalli testaamisvaiheesta eroaa opetusvaiheesta vain viimeisten lohkojen osalta. Testaamisessa sovitetun mallin ja puhesignaalista laskettujen piirteiden avulla lasketaan äänitteelle sujuvuusarvio.

3.4.1 Lineaarinen regressiomalli

Lineaarisen regressiomallin opettamiseen käytettiin 70 % aineistosta. Loppua 30 % datasta käytettiin mallin testaamiseen. Opettamisessa käytettiin kaikkia 16 piirrettä. Opetuksen kohteena (engl. *target*) käytettiin kaikista neljästä arviointikategoriasta yhteenlaskettua pisteytystä, jolloin saatiin uudet arvot 4 ja 20 välille. Jako opetus- ja testidataan tehtiin satunnaisesti. Testi toistettiin sata kertaa eri jaotteluilla opetus- ja testidataan.

Lineaarisen regressiomallin avulla testattiin erikseen myös kaikkien piirteiden kombinaatiot kohteen ollessa kaikista sujuvuuden arviointikriteereistä yhteenlaskettu arvo. Testin tarkoituksena oli löytää

sujuvuuden yhteenlaskettua arviota parhaiten kuvaavat piirteet. Opetusjoukkona käytettiin 70 % kaikesta datasta. Loppua 30 % datasta käytettiin testijoukkona. Jako opetus- ja testijoukkoihin tehtiin satunnaisesti.

3.4.2 Muut tilastolliset mallit

Lineaarisen regression ohella myös muiden suosittujen mallien suorituskkyä testattiin kaikilla piirrekombinaatioilla. Testattuihin malleihin kuuluu lineaarinen regressiomalli, logistinen regressiomalli, lähinaapurinmenetelmä, päätöspuu, satunnaismetsä, perseptroniverkko (engl. *multilayer perceptron*) ja Adaboost. Mallien opettamiseen käytettiin 70 % datasta ja testaamiseen käytettiin loput 30 % datasta. Jako tehtiin satunnaisesti, ja samaa jakoa käytettiin kaikkien mallien opettamiseen ja testaamiseen. Kohteena käytettiin arviointikriteerien yhteenlaskettua arvoa.

Kaikkien mallien testaamisessa käytettiin scikit-learn -kirjaston valmiita malleja. Logistisessa regressiomallissa käytettiin L2-tappiofunktiota (engl. *loss function*). Lähinaapurimenetelmässä käytettiin 5 lähintä naapuria, joiden vaikutus määräytyi niiden euklidisen etäisyyden mukaan. Satunnaismetsämalleissa käytettiin 100 puuta. Perseptroniverkossa käytettiin kolmea kerrosta, joissa oli 16, 100 ja 17 neuronua. Perseptroniverkon iterointirajana oli 100 iteraatiota.

3.5 Piirteiden valitseminen

Jokaista arviointikriteeriä (painotus, sujuvuus, tahti ja tunneilmaisuus) tutkittiin myös erillään toistaan. Arviointikriteerien erottamisen tavoitteena on parantaa arvioinnin tarkkuutta. Vaikka arviointikriteerit ovat toisistaan riippuvia, voidaan olettaa, että tietyt piirteet mittaavat paremmin vain osaa arviointikriteereistä.

Piirteiden vaikutusta eri arviointikriteereihin testattiin lineaarisen regression avulla. Jokaista arviointikriteeriä parhaiten kuvaavat piirteet selvitettiin luvussa 3.4.1 esitellyn piirteiden vertailun tapaan.

3.6 Automaattisen järjestelmän arviointi

Mallien tuottamia arviointituloksia arvioitiin absoluuttisen poikkeaman (engl. *mean absolute error*, *MAE*)

$$MAE = \frac{\sum_{i=1}^n |y_i - x_i|}{n}$$

ja keskineliövirheen neliöjuuren (engl. *root mean squared error*, *RMSE*)

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - x_i)^2}$$

avulla.

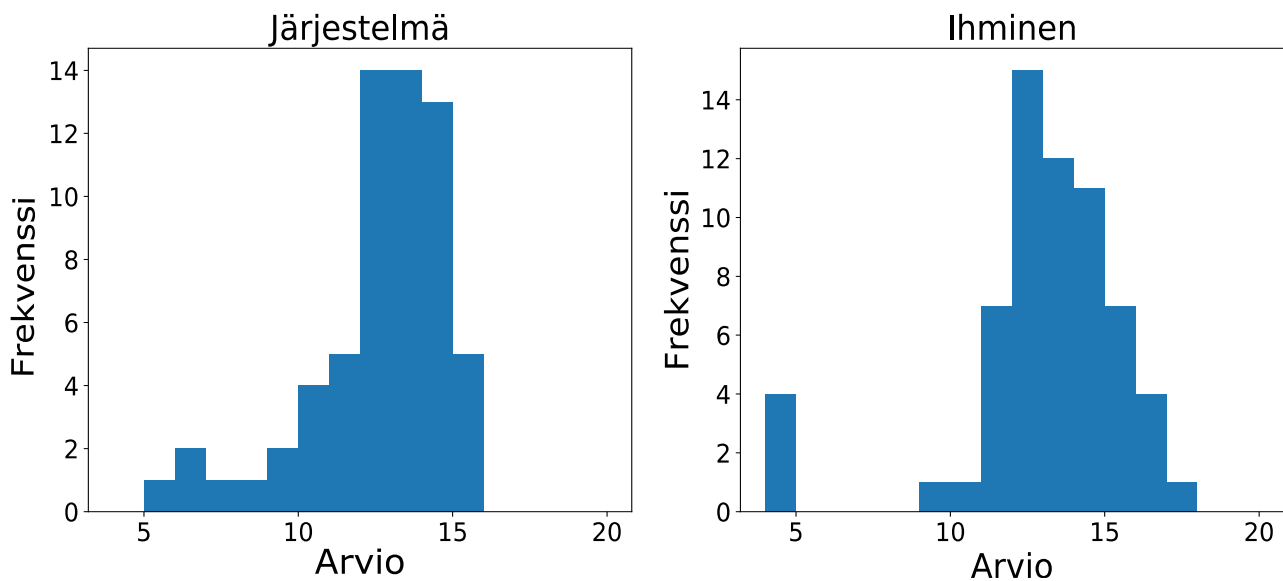
Näiden lisäksi malleille laskettiin R^2 -selitysaste scikit-learn -kirjaston funktiolla `r2_score`. Selitysaste kuvaa kuinka hyvin mallin tuottamat arviot vastaavat todellisia arvoja. Selitysaste tyypillisesti saa arvoja väliltä $[0, 1]$. Mitä lähempänä arvoa 1 selitysaste on, sitä tarkempi malli on. Selitysaste voi olla myös negatiivinen, jolloin malli selittää datan varianssia huonommin kuin datan keskiarvo [9].

Osassa testeistä arviointitulosten tarkkuutta mitattiin myös vertaamalla niitä suoraan ihmisen tekemiin arvioihin. Vertailu toteutettiin laskemalla äänitteiden lukumäärä, joille järjestelmä antoi saman arvion kuin ihminen. Lukumäärä jaettiin kaikkien arvioitujen äänitteiden lukumäärällä, ja muutettiin prosenttiluvuksi.

4. TULOKSET

Automaattisen arvioinnin vertailukohtana voidaan käyttää satunnaisarpojan tuottamaa tulosta. Tulos kuvastaa tarkkuutta, joka saavutetaan silloin kun mallilla ei ole mitään tietoa juuri kyseessä olevasta puhenäytteestä. Arpoja arpoi opetusjoukon sujuvuusarvioiden keskiarvon ja keskihajonnan mukaisesta normaalijakaumasta arvoja, joita verrattiin testijoukon sujuvuusarvioihin. Sadan satunnaisarvon keskimääräinen MAE oli 2,731 ja RMSE 3,541.

Kuvassa 3 on esitetty erään aineistoon sovitetun lineaarisen regressiomallin tuottamien arvioiden histogrammi, sekä vastaavat ihmisen tekemien arvioiden histogrammi. Mallin sovittamiseen on käytetty kaikkia 16 puhesignaalista laskettua piirrettä. Kaikilla piirteillä opetetun mallin absoluuttinen keskipoikkeama oli testeissä välillä 0,995 – 3,247 ja keskineliövirheen neliöjuuri välillä 1,179 – 9,727. Mallin selitysaste oli välillä 0,078 – 0,688.



Kuva 3: Lineaarisen regressiomallin tuottamien arvioiden ja ihmisen tekemien arvioiden histogrammit.

Kymmenestä mallista pienimmän keskineliövirheen neliöjuuren tuottaneen mallin selitysaste oli 0,528. Mallin tuottamien testijoukon arvioiden keskiarvo ja keskihajonta olivat 13,380 ja 1,429. Vastaavasti ihmisen testijoukolle tekemien arvioiden keskiarvo ja keskihajonta olivat 13,143 ja 1,717.

Taulukko 2: Yhteenlaskettujen arvioiden estimointi logistisella regressiolla

Piirteet	MAE	RMSE	R^2
----------	-----	------	-------

TAUK, TAUK _{MAX} , VOK _{KA} , VOK _{MAX} , VOK _{KH} , LOPPU, F0 _{KH} , WPM, WPM _{VAR}	1,386	1,763	0,676
TAUK, TAUK _{MAX} , VOK _{KA} , VOK _{MAX} , VOK _{KH} , VIRK, LOPPU, F0 _{KH} , WPM, WPM _{VAR}	1,376	1,764	0,675
TAUK, TAUK _{MAX} , VOK _{KA} , VOK _{MAX} , VOK _{KH} , KONS _{KH} , LOPPU, F0 _{KH} , WPM, WPM _{VAR}	1,373	1,765	0,675
TAUK, TAUK _{MAX} , TAUK _{KH} , VOK _{KA} , VOK _{MAX} , KONS _{KH} , LOPPU, F0 _{KA} , F0 _{KH} , WPM, WPM _{VAR}	1,369	1,765	0,675
TAUK, TAUK _{MAX} , TAUK _{KH} , VOK _{KA} , VOK _{MAX} , VIRK, LOPPU, F0 _{KA} , F0 _{KH} , WPM, WPM _{VAR}	1,372	1,767	0,674
TAUK, TAUK _{MAX} , TAUK _{KH} , VOK _{KA} , VOK _{MAX} , KONS _{MAX} , KONS _{KH} , LOPPU, F0 _{KA} , F0 _{KH} , WPM, WPM _{VAR}	1,371	1,769	0,674
TAUK, TAUK _{MAX} , TAUK _{KH} , VOK _{KA} , VOK _{MAX} , KONS _{MAX} , KONS _{KH} , LOPPU, F0 _{KH} , WPM, WPM _{VAR}	1,366	1,769	0,673
TAUK, TAUK _{MAX} , TAUK _{KH} , VOK _{KA} , VOK _{MAX} , KONS _{MAX} , LOPPU, F0 _{KH} , WPM, WPM _{VAR}	1,374	1,770	0,673
TAUK, TAUK _{MAX} , TAUK _{KH} , VOK _{KA} , VOK _{MAX} , LOPPU, F0 _{KA} , F0 _{KH} , WPM, WPM _{VAR}	1,374	1,771	0,673
TAUK, TAUK _{MAX} , TAUK _{KH} , VOK _{KA} , VOK _{MAX} , KONS _{KH} , VIRK, LOPPU, F0 _{KA} , F0 _{KH} , WPM, WPM _{VAR}	1,378	1,772	0,672
WPM	1,633	2,106	0,537
Kaikki piirteet	2,192	5,954	-2,700

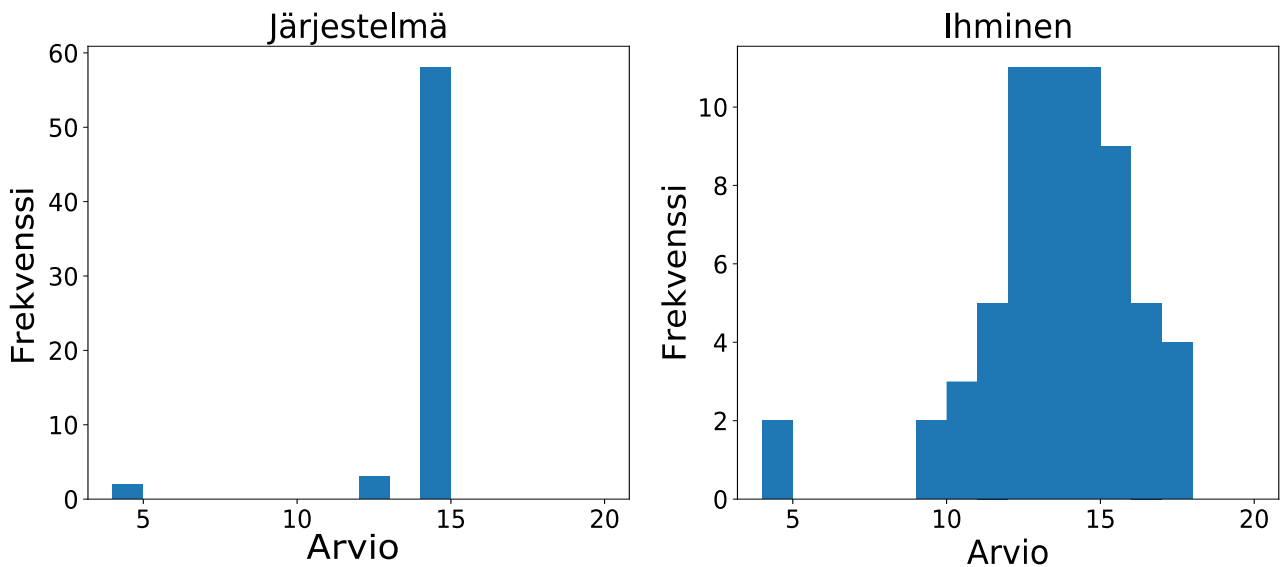
Taulukossa 2 on esiteltynä kombinaatiot, jotka tuottivat parhaiten ihmisen antamaa arviota vastaavan arvion. Näiden kombinaatioiden lisäksi taulukossa on vertailun vuoksi myös pelkällä WPM:llä opetettu malli sekä kaikilla piirteillä opetettu malli. Piirteet TAUK, TAUK_{MAX}, VOK_{KA}, VOK_{MAX}, VOK_{KH}, LOPPU, F0_{KH}, WPM ja WPM_{VAR} esiintyivät kaikissa kymmenessä kombinaatiossa ja piirteet KONS_{KH} ja F0_{KA} esiintyvät puolessa niistä. Piirteet TAUK_{KA}, TAUK_{KH} ja KONS_{KA} eivät esiintyneet yhdessäkään hyvin toimivassa mallissa.

Kaikilla piirteillä sovitettun mallin arviointitulos oli tosi heikko. Tämä johtui heikosti puheentunnistimella tunnistetuiden äänitteiden olemisesta testijoukossa. Näiden äänitteiden kohdalla tietyt piirteet saivat hyvin poikkeavia arvoja. Erityisesti piirteet TAUK_{KA}, ja TAUK_{KH} saivat hyvin erilaisia arvoja poikkeavien tunnistuksien takia.

Taulukko 3: Yhteenlaskettujen arvioiden estimointi eri malleilla. Kunkin mallin tulos vastaa kyseisellä mallilla parhaiten toiminutta piirrekombinaatiota.

Malli	Piirteet	MAE	RMSE	R ²
LinearRegression	TAUK, VOK _{KA} , VOK _{MAX} , VOK _{KH} , LOPPU, F0 _{KA} , F0 _{KH} , WPM	1,048	1,187	0,716
LogisticRegression	F0 _{KA} , F0 _{KH} , WPM	0,937	1,189	0,715

<i>KNeighbors</i>	TAUK _{MAX} , VOK _{MAX} , WPM	1,413	1,821	0,331
<i>DecisionTree</i>	TAUK, VOK _{KA} , KONS _{KH} , LOPPU, F0 _{KA} , F0 _{KH} , WPM, WPM _{VAR}	1,111	1,480	0,558
<i>RandomForest</i>	TAUK, VOK _{MAX} , VOK _{KA} , F0 _{KH} , WPM	1,032	1,303	0,657
<i>MLP</i>	TAUK, TAUK _{MAX} , VOK _{MAX} , VOK _{KA} , LOPPU, WPM	1,079	1,380	0,616
<i>Adaboost</i>	WPM	1,238	1,512	0,539



Kuva 4: AdaBoost-mallin tuottamat arviot puheen sujuvuudelle

Taulukossa 3 on esitelty jokaisella testatulla mallilla tai metodilla tuotettu paras arviointitulos sekä opettamiseen käytetty piirrekombinaatio. Vaikka logistisella regressiomallilla ja AdaBoost-mallilla RMSE on pieni, histogrammia tarkastelemalla nähdään, ettei malli ole yleistynyt dataan hyvin. Kuvassa 4 on AdaBoost-mallin tuottamien arvioiden histogrammi. Matalaa keskineliövirhettä selittää yhteisarvion pieni hajonta. Koko aineiston yhteisarvioiden keskihajonta on 2,47. Keskihajonnan ollessa pieni, voi pelkästään odotusarvolla käyttämällä päästä matalaan keskineliövirheeseen. Muut mallit yleistyivät paremmin opetusdataan. Vertailun perusteella lineaarinen regressiomalli, satunnaismetsämenetelmä ja perseptroniverkko ovat otollisimmat tilastolliset mallit.

Taulukko 4: Painotusta kuvaavimmat piirteet

Piirteet	MAE	RMSE	R^2	Oikeita arvioita
TAUK, TAUK _{KA} , TAUK _{MAX} , VOK _{KA} , VOK _{MAX} , VOK _{KH} , KONS _{KA} , KONS _{MAX} , KONS _{KH} , LOPPU, WPM _{VAR}	0,487	0,582	0,446	58,7 %

TAUK, TAUK _{KA} , TAUK _{MAX} , VOK _{KA} , VOK _{MAX} , VOK _{KH} , KONS _{KA} , KONS _{MAX} , KONS _{KH} , LOPPU, F0 _{KA} , WPM _{VAR}	0,488	0,583	0,446	58,7 %
TAUK, TAUK _{KA} , TAUK _{MAX} , VOK _{KA} , VOK _{MAX} , VOK _{KH} , KONS _{KA} , KONS _{MAX} , KONS _{KH} , LOPPU, F0 _{KH} , WPM _{VAR}	0,488	0,584	0,443	57,1 %
TAUK, TAUK _{KA} , TAUK _{MAX} , VOK _{KA} , VOK _{MAX} , VOK _{KH} , LOPPU, WPM _{VAR}	0,498	0,586	0,439	60,3 %
TAUK, TAUK _{KA} , TAUK _{MAX} , VOK _{KA} , VOK _{MAX} , VOK _{KH} , LOPPU, F0 _{KA} , WPM _{VAR}	0,499	0,586	0,439	60,3 %
TAUK, TAUK _{KA} , TAUK _{MAX} , VOK _{KA} , VOK _{MAX} , VOK _{KH} , LOPPU, F0 _{KH} , WPM _{VAR}	0,499	0,587	0,439	58,7 %
TAUK, TAUK _{MAX} , VOK _{KA} , VOK _{MAX} , VOK _{KH} , KONS _{KA} , KONS _{MAX} , KONS _{KH} , LOPPU, F0 _{KA}	0,492	0,588	0,436	57,1 %
TAUK, TAUK _{MAX} , VOK _{KA} , VOK _{MAX} , VOK _{KH} , KONS _{KA} , KONS _{MAX} , KONS _{KH} , LOPPU	0,492	0,588	0,436	57,1 %
TAUK, TAUK _{KA} , TAUK _{MAX} , VOK _{KA} , VOK _{MAX} , VOK _{KH} , KONS _{KA} , KONS _{MAX} , KONS _{KH} , VIRK, LOPPU, WPM _{VAR}	0,501	0,589	0,434	55,6 %
TAUK, TAUK _{KA} , VOK _{KA} , VOK _{MAX} , VOK _{KH} , KONS _{KA} , KONS _{MAX} , KONS _{KH} , LOPPU, F0 _{KH}	0,491	0,589	0,433	63,5 %
WPM	0,563	0,690	0,224	50,8 %

Taulukko 5: Sujuvuutta kuvaavimmat piirteet

Piirteet	MAE	RMSE	R ²	Oikeita arvioita
TAUK, TAUK _{MAX} , VOK _{KA} , VOK _{MAX} , VOK _{KH} , VIRK, LOPPU, F0 _{KA} , WPM	0,351	0,469	0,658	73,0 %
TAUK, TAUK _{MAX} , TAUK _{KH} , VOK _{KA} , VOK _{MAX} , VOK _{KH} , VIRK, LOPPU, F0 _{KA} , WPM	0,350	0,468	0,658	73,0 %
TAUK, TAUK _{MAX} , VOK _{KA} , VOK _{MAX} , VOK _{KH} , KONS _{KA} , VIRK, LOPPU, F0 _{KA} , WPM	0,352	0,468	0,658	74,6 %
TAUK, TAUK _{KA} , TAUK _{KH} , VOK _{KA} , VOK _{MAX} , VOK _{KH} , VIRK, LOPPU, F0 _{KA} , WPM	0,352	0,468	0,657	73,0 %
TAUK, TAUK _{MAX} , TAUK _{KH} , VOK _{KA} , VOK _{MAX} , VOK _{KH} , KONS _{KA} , VIRK, LOPPU, F0 _{KA} , WPM	0,351	0,468	0,657	74,6 %
TAUK, TAUK _{MAX} , VOK _{KA} , VOK _{MAX} , VOK _{KH} , LOPPU, F0 _{KA} , WPM	0,353	0,469	0,657	73,0 %
TAUK, TAUK _{MAX} , VOK _{KA} , VOK _{MAX} , VOK _{KH} , VIRK, LOPPU, F0 _{KH} , WPM	0,355	0,469	0,657	74,6 %

TAUK, TAU _{KMAX} , TAU _{KH} , VOK _{KA} , VOK _{MAX} , VOK _{KH} , LOPPU, F0 _{KA} , WPM	0,352	0,469	0,657	73,0 %
TAUK, TAU _{KA} , TAU _{KMAX} , TAU _{KH} , VOK _{KA} , VOK _{MAX} , VOK _{KH} , VIRK, LOPPU, F0 _{KA} , WPM	0,353	0,469	0,657	73,0 %
TAUK, TAU _{KA} , TAU _{KH} , VOK _{KA} , VOK _{MAX} , VOK _{KH} , KONS _{KA} , VIRK, LOPPU, F0 _{KA} , WPM	0,353	0,469	0,656	73,0 %
WPM	0,419	0,547	0,532	71,4 %

Taulukko 6: Tahtia kuvaavimmat piirteet

Piirteet	MAE	RMSE	R ²	Oikeita arvioita
TAUK, TAU _{KA} , TAU _{KMAX} , TAU _{KH} , VOK _{KA} , VOK _{MAX} , VOK _{KH} , KONS _{KA} , VIRK, LOPPU, F0 _{KA} , F0 _{KH} , WPM, WPM _{VAR}	0,413	0,518	0,697	63,5 %
TAUK, TAU _{KA} , TAU _{KH} , VOK _{KA} , VOK _{MAX} , VOK _{KH} , KONS _{KA} , VIRK, LOPPU, F0 _{KA} , F0 _{KH} , WPM, WPM _{VAR}	0,415	0,519	0,695	63,5 %
TAUK, TAU _{KA} , TAU _{KMAX} , TAU _{KH} , VOK _{MAX} , VOK _{KH} , KONS _{KA} , VIRK, LOPPU, F0 _{KA} , F0 _{KH} , WPM, WPM _{VAR}	0,415	0,519	0,695	65,1 %
TAUK, TAU _{KA} , TAU _{KH} , VOK _{MAX} , VOK _{KH} , KONS _{KA} , VIRK, LOPPU, F0 _{KA} , F0 _{KH} , WPM, WPM _{VAR}	0,416	0,520	0,695	65,1 %
TAUK, TAU _{KA} , TAU _{KH} , VOK _{MAX} , VOK _{KH} , KONS _{KA} , VIRK, LOPPU, WPM, 16	0,420	0,521	0,693	63,5 %
TAUK, TAU _{KA} , TAU _{KMAX} , TAU _{KH} , VOK _{MAX} , VOK _{KH} , KONS _{KA} , VIRK, LOPPU, WPM, 16	0,421	0,521	0,693	63,5 %
TAUK, TAU _{KA} , TAU _{KH} , VOK _{KA} , VOK _{MAX} , VOK _{KH} , KONS _{KA} , VIRK, LOPPU, WPM, 16	0,419	0,522	0,692	65,1 %
TAUK, TAU _{KA} , TAU _{KH} , VOK _{KA} , VOK _{MAX} , VOK _{KH} , KONS _{KA} , LOPPU, F0 _{KA} , F0 _{KH} , WPM, WPM _{VAR}	0,418	0,522	0,692	63,5 %
TAUK, TAU _{KA} , TAU _{KH} , VOK _{KA} , VOK _{MAX} , VOK _{KH} , KONS _{KA} , VIRK, LOPPU, F0 _{KH} , WPM, WPM _{VAR}	0,419	0,522	0,692	65,1 %
TAUK, TAU _{KA} , TAU _{KMAX} , TAU _{KH} , VOK _{KA} , VOK _{MAX} , VOK _{KH} , KONS _{KA} , VIRK, LOPPU, WPM, WPM _{VAR}	0,423	0,522	0,692	65,1 %
WPM	0,444	0,574	0,628	58,7 %

Taulukko 7: Tunneilmaisua kuvaavimmat piirteet

Piirteet	MAE	RMSE	R ²	Oikeita arvioita
TAUK, TAU _{KMAX} , VOK _{KA} , VOK _{KH} , KONS _{KA} , VIRK, LOPPU, F0 _{KA} , F0 _{KH}	0,510	0,618	0,280	50,8 %

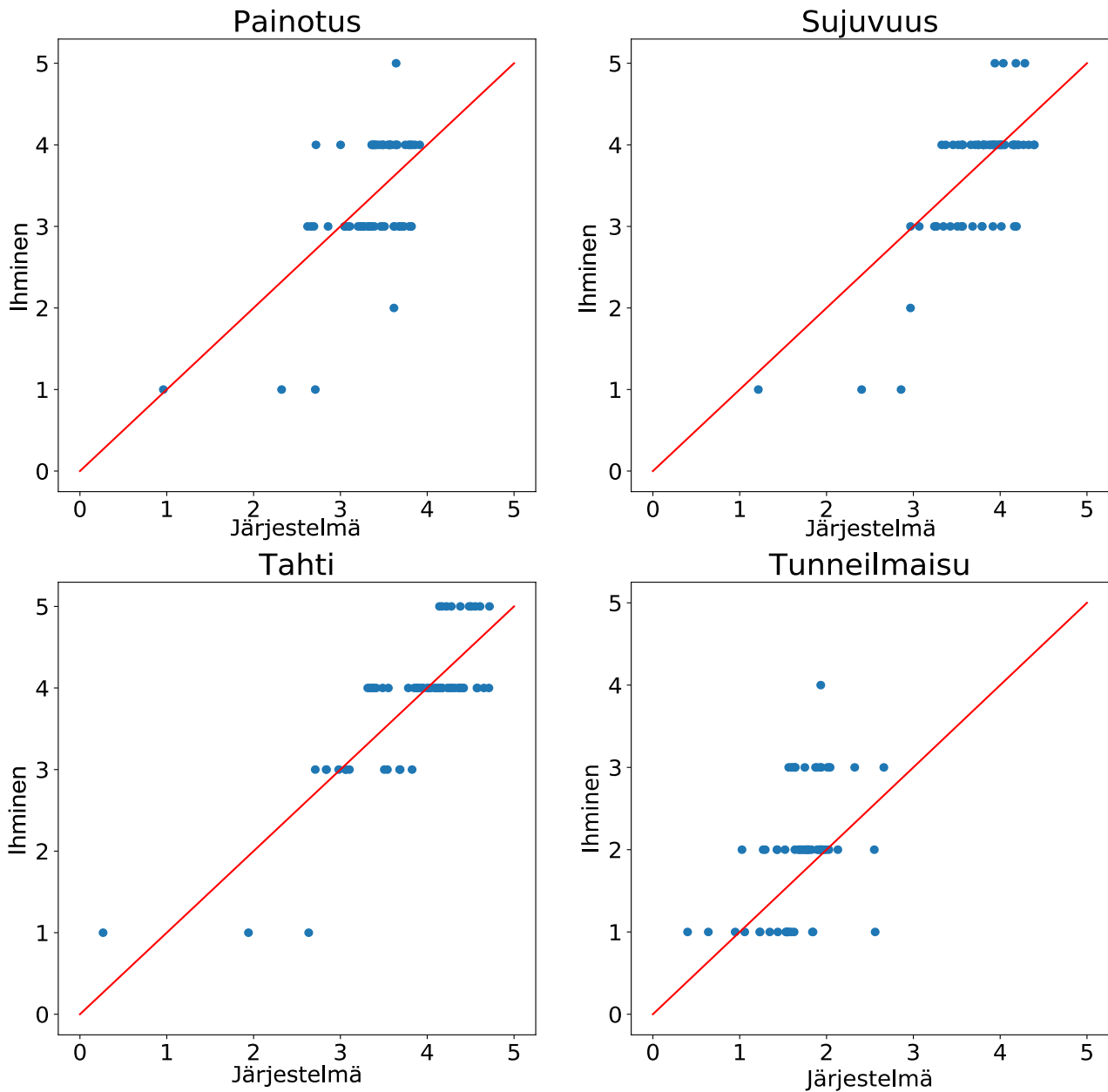
TAUK, TAUK _{MAX} , VOK _{KA} , VOK _{KH} , KONS _{KA} , LOPPU, F0 _{KA} , F0 _{KH}	0,510	0,618	0,279	50,8 %
TAUK, TAUK _{MAX} , VOK _{KA} , VOK _{KH} , KONS _{KA} , VIRK, F0 _{KA} , F0 _{KH}	0,510	0,619	0,278	49,2 %
TAUK, TAUK _{MAX} , VOK _{KA} , VOK _{KH} , KONS _{KA} , F0 _{KA} , F0 _{KH}	0,510	0,620	0,277	50,8 %
TAUK, TAUK _{MAX} , VOK _{KA} , VOK _{MAX} , VOK _{KH} , KONS _{KA} , VIRK, LOPPU, F0 _{KA} , F0 _{KH}	0,520	0,620	0,276	49,2 %
TAUK, TAUK _{MAX} , VOK _{MAX} , VOK _{KH} , KONS _{KA} , VIRK, LOPPU, F0 _{KA} , F0 _{KH} , WPM, WPM _{VAR}	0,513	0,620	0,275	49,2 %
TAUK, TAUK _{MAX} , VOK _{MAX} , VOK _{KH} , KONS _{KA} , LOPPU, F0 _{KA} , F0 _{KH} , WPM, WPM _{VAR}	0,513	0,621	0,274	47,6 %
TAUK, TAUK _{MAX} , VOK _{MAX} , VOK _{KH} , KONS _{KA} , VIRK, F0 _{KA} , F0 _{KH} , WPM, WPM _{VAR}	0,514	0,621	0,274	49,2 %
TAUK, TAUK _{MAX} , VOK _{KA} , VOK _{MAX} , VOK _{KH} , KONS _{KA} , LOPPU, F0 _{KA} , F0 _{KH}	0,520	0,621	0,273	49,2 %
TAUK, TAUK _{MAX} , VOK _{MAX} , VOK _{KH} , KONS _{KA} , F0 _{KA} , F0 _{KH} , WPM, WPM _{VAR}	0,515	0,622	0,272	47,6 %
WPM	0,570	0,704	0,067	46,0 %

Arviointikriteereistä sujuvuuden arvioiminen onnistuu parhaiten ja tunneilmaisun arviointi huonoiten. Sujuvuuden arvioinnissa päästiin valikoituja piirteitä käyttäen yli 70 % tarkkuuteen. Aiemmin toteutettu englanninkielisen äänen lukemisen arviointiin tarkoitettu FLORA-järjestelmä pääsi 76 % tarkkuuteen neliasteisen sujuvuuden arvioinnissa [6].

Taulukkojen 4–7 esiintymisfrekvenssin perusteella painotuksen arvioimiseen vaikuttavat piirteet TAUK, VOK_{KA}, VOK_{MAX}, VOK_{KH} ja LOPPU eniten. Sujuvuuden arviointiin vaikuttavat puolestaan piirteet TAUK, VOK_{KA}, VOK_{MAX}, VOK_{KH}, KONS_{KH} ja WPM. Tahdin arvioinnissa tärkeimmät piirteet ovat TAUK, TAUK_{KA}, TAUK_{KH}, VOK_{KH}, KONS_{KA}, LOPPU, WPM ja WPM_{KA}, ja tunneilmaisun arvioinnissa piirteet TAUK, TAUK_{KA}, VOK_{KH}, KONS_{KA}, F0_{KA} ja F0_{KH}.

Kaikki piirteet esiintyvät ainakin kerran jonkun arviointikriteerin kuvaavimmissa piirteissä. Tämän perusteella niitä voidaan pitää onnistuneina. Taukojen määrä ja vokaalien pituuden keskihajonta vaikuttavat merkittävästi kaikkiin arviointikriteereihin. Äänen perussävelen vaihtelua kuvaavat piirteet vaikuttivat eniten tunneilmaisuuksiin. Tärkeä huomio on piirteen LOPPU korkea frekvenssi taulukoissa. Piirre LOPPU kuvaa ASR:n tunnistusvirheiden määrää, joita esiintyy lukemisen ollessa toistavaa ja/tai katkonaista. Siten toistojen määrää ja sanojen sisäisiä taukoja kuvaavat piirteet voivat olla hyvin kuvaavia.

Kuvassa 5 on esitelty automaattisen arvioinnin tuloksia ihmisen tekemien arvioiden suhteen. Pystyakselilla on ihmisen tekemät arviot ja vaaka-akselilla automaattisesti luodut arviot. Parhaassa mahdollisessa tilanteessa kaikki pisteet olisivat viivan kohdalla, jolloin automaattinen järjestelmä tuottaisi saman arvion kuin ihminen.



Kuva 5: Automaattiset arviointitulokset ihmisen antaman arvioinnin suhteen

Kuten luvussa 2.2 todettiin, lasten puheen tunnistaminen on vaikeampaa kuin aikuisten. Työssä käytetty AaltoASR on opetettu aikuisten ihmisten puheella, joten lasten puheen tunnistustarkkuus ei

tuottanut toivottua tulosta. TextGrid-tiedostojen aikaleimat päättyivät joissain tapauksissa huomattavasti aikaisemmin, kuin äänitteen todelliset puhekohdat päättyivät. Etenkin äänitteet, joissa lukeminen takelteli, ASR:n tunnistus tulos kärsi valtavasti. Aikaleimojen epätarkkuus luultavasti heikensi järjestelmän toimintaa, koska se väärensi joidenkin piirteiden arvoja. Ratkaisuna tähän on käyttää lasten puheella opetettua puheentunnistinta tai tähän tehtävään räätälöityä puheentunnistinta. Esimerkiksi NWEA:n lukupuheen sujuvuuden arviointiin kehittämä työkalu [25] käyttää erityistä EduSpeak®-puheentunnistinta [12], joka on tarkoitettu 4–8-vuotiaiden lasten puheentunnistamiseen.

Työssä käytetty aineisto on suhteellisen pieni. Äänitteiden kokonaiskesto on noin 6,5 tuntia, ja lukijoiden määrä on 108. Bolaños et al. julkaisussa [6] käytetty aineisto sisälsi 13 tuntia äänitteitä yli 300 oppilaan puhumana. Suuri lukijoiden määrä lisää varianssia datassa, joka voi parantaa mallin soveltamisen tarkkuutta.

Usein lukemisen sujuvuuden arviointikriteerinä käytetty WPM osoitti olevan yhteydessä kaikkiin arviointikriteereihin. Prosodisien piirteiden käyttäminen WPM:n ohella paransi arvioinnin tarkkuutta kaikilla kriteereillä. Myös kriteerikohtaisten piirteiden käyttäminen paransi arvioinnin tarkkuutta. Voidaan olettaa, että kriteerejä hyvin kuvaavien piirteiden lisääminen tunnistukseen parantaisi tulosta entistään. Lisäksi räätälöidyn puheentunnistimen käyttäminen mahdollistaisi monipuolisempien piirteiden laskemisen.

5. YHTEENVETO

Työn tarkoitus oli luoda järjestelmä, joka arvioi automaattisesti äänen lukemisen sujuvuutta. Sujuvuuden arviointia tutkittiin neljän eri sujuvuuden kriteerin kautta. Sujuvuuden arviointiin käytettiin puheentunnistimen ja puheen perussävelentunnistimen avulla tuotettuja piirteitä. Piirteiden tarkoitus oli mitata ääneen lukemisesta sujuvuutta tai takkuilevuutta kuvaavia ominaisuuksia. Työssä käytettiin 16 puhesignaalista laskettua piirrettä, joihin kuului muun muassa lauseiden välisten taukojen pituus ja luettujen sanojen määrä minuutissa. Ääneen lukemisen sujuvuutta arvioitiin sovittamalla äänitteestä lasketut piirteet ja ihmisen äänitteille antamat lineaarisen regressiomallin avulla.

Sujuvuuden eri kriteerien arvioinnin kannalta tärkeät piirteet löydettiin kaikista käytetyistä piirteistä. Piirteiden valikoiminen paransi arviointia. Tärkeiden piirteiden pohjalta voidaan kehittää myöhemmin uusia kriteerejä kuvaavia piirteitä.

Työssä tuotettua järjestelmää voidaan pitää onnistuneena. Järjestelmällä päästään sujuvuuden arvioinnissa 70 % tarkkuuteen, joka on lähes yhtä hyvä tarkkuus kuin aikaisemmin toteutetussa englanninkieliltä arvioivassa järjestelmässä.

LÄHTEET

- [1] Aalto-ASR – automaattinen puheentunnistin, Kielipankki, verkkosivu. Saatavissa (viitattu 24.11.2019):
<https://www.kielipankki.fi/tuki/aalto-asr-automaattinen-puheentunnistin/>
- [2] M. Airaksinen, L. Juvela, P. Alku, O. Räsänen, Data Augmentation Strategies for Neural Network F0 Estimation, ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brighton, United Kingdom, United Kingdom, May 12–17 2019. IEEE, pp. 6485–6489.
- [3] R.G. Benjamin, P.J. Schwanenflugel, E.B. Meisinger, C. Groff, M.R. Kuhn, L. Steiner, A Spectrographically Grounded Scale for Evaluating Reading Expressiveness, Reading Research Quarterly, Vol. 48, No. 2, 2013, pp. 105–133.
- [4] R.G. Benjamin, P.J. Schwanenflugel, Text Complexity and Oral Reading Prosody in Young Readers, Reading Research Quarterly, Vol. 45, Iss. 4, 2010, pp. 388–404.
- [5] J. Bernstein, A. Van Moere, J. Cheng, Validating automated speaking tests, Language Testing, Vol. 27, Iss. 3, 2010, pp. 355–377.
- [6] D. Bolaños, R.A. Cole, W.H. Ward, G.A. Tindal, P.J. Schwanenflugel, M.R. Kuhn, Automatic assessment of expressive oral reading, Speech Communication, Vol 55, Iss. 2, 2013, pp. 221–236.
- [7] J. Cheng, J. Shen, Towards Accurate Recognition for Children’s Oral Reading Fluency, 2010 IEEE Spoken Language Technology Workshop, 2010, pp. 103–108.
- [8] M.M. Clay, R.H. Imlach, Juncture, pitch, and stress as reading behavior variables, Journal of Verbal Learning and Verbal Behavior, Vol. 10, No. 2, 1971, pp. 133–139.
- [9] A. Coster, Goodness-of-Fit Statistics, UNSW Sydney, verkkosivu. Saatavissa (viitattu 1.12.2019):
<https://web.maths.unsw.edu.au/~adelle/Garvan/Assays/GoodnessOfFit.html>
- [10] T. Elliott, The State of the Octoverse: machine learning, The GitHub Blog, verkkosivu. Saatavissa (viitattu 29.11.2019):
<https://github.blog/2019-01-24-the-state-of-the-octoverse-machine-learning/>
- [11] J. Hasbrouck, G.A. Tindal, Oral reading fluency norms: A valuable assessment tool for reading teachers, The Reading Teacher, Vol. 59, Iss. 7, 2006, pp. 636–644.
- [12] F. Horacio, H. Bratt, R. Rossier, G. Rao, S.E. Venkata, EduSpeak®: A speech recognition and pronunciation scoring toolkit for computer-aided language learning applications, Language Testing, Vol. 27, Iss. 3, 2010, pp. 401–418.
- [13] A. Huhtala, V. Ravolainen, Lukemisen ja kirjoittamisen testejä, Finra ry, Suomi, 2010, 134 s.
- [14] T. Häyrynen, S. Serenius-Sirve, M. Korkman, Lukulasse, Psykologinen Kustannus Oy, Suomi, 1999, 112 s.

- [15] A. Kajamies, E. Poskiparta, T. Annevirta, M. Dufva, M. Vauras, YTTE – Luetun ja kuullun ymmärtämisen ja lukemisen sujuvuuden arviointi, Turun yliopisto, Oppimistutkimuksen keskus, 2003, 179 s.
- [16] S. Kee, A. Potamianos, S. Narayanan, Acoustics of children's speech: Developmental changes of temporal and spectral parameters, Journal of Acoustic Society of America, Vol. 105, No. 3, 1999, pp. 1455–1468.
- [17] M. Kurimo et al. Aalto University Automatic Speech Recognition System. Saatavilla (viitattu 4.12.2019): <http://urn.fi/urn:nbn:fi:lb-2014091904>
- [18] S. Lemmetty, Review of Speech Synthesis Technology, Helsinki University of Technology, 1999, 76 p. Saatavissa (viitattu 27.11.2019): http://research.spa.aalto.fi/publications/theses/lemmetty_mst/index.html
- [19] M.-K. Lerkkanen, A.-M. Poikkeus, R. Ketonen, Armi 1. Luku- ja kirjoitustaidon arviointimateriaali 1. luokalle, WSOY Oppimateriaalit, Suomi, 2006, 82 s.
- [20] M.-K. Lerkkanen, A.-M. Poikkeus, R. Ketonen, Armi 2. Luku- ja kirjoitustaidon arviointimateriaali 2. luokalle, WSOY Oppimateriaalit, Suomi, 2008, 96 s.
- [21] Q. Li, M. Russel, An Analysis of the causes of increased error rates in children's speech recognition, ICSLP 2002, 7th International Conference on Spoken Language Processing, Denver, Colorado, USA, September 16–20, 2002. pp. 2337–2340.
- [22] P. Lin, Y. Su, Y. Huang, Evaluating reading fluency behavior via reading rates of elementary school students reading e-books, Computers in Human Behavior, Vol. 100, 2019, pp 258–265.
- [23] A. Loukina, N. Madnani, B. Beigman Klebanov, A. Misra, G. Angelov, O. Todric, Evaluating on-device ASR on Field Recordings from an Interactive Reading Companion, 2018 IEEE Spoken Language Technology Workshop, 2018, pp. 964–970.
- [24] J. Miller, P.J. Schwanenflugel, Prosody of Syntactically Complex Sentences in the Oral Reading of Young Children, Journal of Educational Psychology, Vol. 94, Iss. 4, 2006, pp. 839–853.
- [25] NWEA Launches MAP Reading Fluency, First and Only K-3 Oral Reading Assessment Using Speech Recognition, Automatic Scoring, Computer Adaptive Technology, Cision PR Newswire, verkkosivu. Saatavissa (viitattu 3.12.2019): <https://www.prnewswire.com/news-releases/nwea-launches-map-reading-fluency-first-and-only-k-3-oral-reading-assessment-using-speech-recognition-automatic-scoring-computer-adaptive-technology-300586162.html>
- [26] PIAAC 2012. Kansainvälisen aikuistutkimuksen ensituloksia, Opetus- ja kulttuuriministeriö, Opetus- ja kulttuuriministeriön julkaisuja 2013:19. Saatavissa (viitattu 23.11.2019): <http://urn.fi/URN:ISBN:978-952-263-237-1>
- [27] G.S. Pinnell, J.J. Pikulksi, K.K. Wixson, J.R. Campbell, P.B. Gough, A.S. Beatty, Listening to Children Read Aloud: data from NAEP's integrated reading performance record (IRPR) at grade 4, National Center for Education Statistics, Report No. 23-FR-04, 1995, 98 p. Saatavissa (viitattu 26.11.2019): <https://catalog.hathitrust.org/Record/003007646>
- [28] PISA 15 Ensituloksia, Opetus- ja kulttuuriministeriö, Opetus- ja kulttuuriministeriön julkaisuja 2016:41. Saatavissa (viitattu 23.11.2019): <http://urn.fi/URN:ISBN:978-952-263-436-8>

- [29] Pitch Estimation Network, GitHub, verkkosivu. Saatavissa (viitattu 24.11.2019): <https://github.com/mairaksi/PIENet>
- [30] E. Poskiparta, P. Niemi, J. Lepola, Diagnostiset testit: 1, Lukeminen ja kirjoittaminen, Oppimistutkimuksen keskus, Turun Yliopisto, Suomi, 1994, 48 s.
- [31] T.V. Rasinski, Assessing Reading Fluency, Pacific Resources for Education and Learning, U.S. Department of Education, USA, 2004, 25 p.
- [32] T. Rasinski, A. Rikli, S. Johnston, Reading Fluency: More Than Automaticity? More Than a Concern for the Primary Grades?, Literacy Research and Instruction, Vol. 48, Iss. 4, 2009, pp. 350–361.
- [33] ReadDrama, Niilo Mäki Instituutti, verkkosivu. Saatavissa (viitattu 26.11.2019): <https://readrama.nmi.fi/>
- [34] Report of the National Reading Panel: Teaching Children to Read, National Reading Panel, 2000, 435 p. Saatavissa (viitattu 26.11.2019): <https://www.nichd.nih.gov/publications/pubs/nrp/smallbook>
- [35] K. Sabu, P. Rao, Automatic assessment of children's oral reading using speech recognition and prosody modeling, CSI Transactions on ICT, Vol 6, Iss. 2, pp. 221–225.
- [36] M. Schall, C.H. Skinner, S. Cazzell, D. Ciancio, J. Ruddy, K. Thompson, Extending Research on Oral Reading Fluency Measures, Reading Speed, and Comprehension, Contemporary School Psychology, Vol. 20, Iss. 3, 2016, pp. 262–269.
- [37] Scikit-learn, verkkosivu. Saatavissa (viitattu 29.11.2019): <https://scikit-learn.org/stable/index.html>
- [38] Sujuvuuden arviointi, Lukimat, verkkosivu. Saatavissa (viitattu 24.11.2019): <http://www.lukimat.fi/lukeminen/tietopalvelu/arviointi/lukemisen-sujuvuuden-arviointi>
- [39] S.W. Valencia, A.T. Smith, A.M. Reece, M. Li, K.K. Wixson, H. Newman, Oral Reading Fluency Assessment: Issues of Construct, Criterion, and Consequential Validity, Reading Research Quarterly, Vol. 45, No. 3, 2010, pp. 270–291.
- [40] A. Viterbi, Error bounds for convolutional codes and an asymptotically optimum decoding algorithm, IEEE Transactions on Information Theory, Vol. 13, Iss. 2, 1967, pp. 260–269.
- [41] J. Zutell, T.V. Rasinski, Training Teachers to Attend to Their Students Oral Reading Fluency, Theory Into Practice, Vol 30, No. 3, 1991, pp. 211–217.

LIITE A: AINEISTOSSA KÄYTETYT TARINAT

Paketti

Aina joskus lapset kohtaavat uusia asioita. Sellaisia, jotka tulevat aikuisten maailmasta. Näin kävi Virpille, kun äiti saapui postilaatikolta ilmoittaen: "Arvaapa mitä! Sinulle on tullut postipaketti." Virpi hämmäntyi saadessaan käteensä paperisen ilmoituksen. "Eihän tämä ole mikään paketti, pelkkä paperilappu", Virpi ihmetteli ääneen. "Mitä minun pitäisi tällä lapulla oikein tehdä?" Äiti selitti, että lappu on vietävä postiin, jossa virkailija antaisi hänelle paketin. Virpiä alkoi jännittämään. Mitä paketissa mahtoi olla, ja kuka sen oli lähettänyt? Postissa oli aikamoinen jono, mikä koetteli Virpin kärsivällisyyttä. Loputtomalta tuntuneen odottelun jälkeen oli viimein hänen vuoronsa. "Minulle on saapunut postipaketti. Tässä on siitä ilmoitus", totesi Virpi topakasti. Virkailija naputteli hetken tietokoneella ja pyysi Virpiltä allekirjoituksen. Sitten hän haki paketin hyllyltä ja ojensi sen Virpille. Kotiin päästyä Virpi avasi heti paketin. Sen sisältä löytyi uusi tablet-tietokone, jonka äiti oli tilannut hänelle salaa. "Hyvää syntymäpäivää pikkuiseni", onnitteli äiti Virpiä. Virpin syntymäpäivä on vasta ensi viikolla, mutta joskus paketit saapuvat perille etuajassa.

Bike

Matilla oli ongelma. Sellainen pieni harmitus, joka tulee mieleen lähes päivittäin. Hei nimittäin osannut ajaa uudella ja hienolla apupyörättömällä polkupyörällään. "Taitaa uusi polkupyörä jäädä käyttämättä, kun en opi ajamaan". huokaisi Matti apeana. Myös isä oli huomannut Matilla olevan vaikeuksia oppia ajamaan polkupyörällään. "Poikani, nyt me ryhdymme harjoittelemaan polkupyörällä ajamista joka päivä!" kajautti isä juhlallisesti. Äitiä isän tomeruus hymyilytti, olihan isällä tapana toisinaan lupaila liikoja. Siitä päivästä lähtien parivaljakko oli kuitenkin tuttu näky kotikadulla. "Yritetäänpä uudestaan, polkaise vain rohkeasti, jotta pääset vauhtiin!" hoki isä. "Eihän tästä tule yhtään mitään. Tyhmää koko pyöräily!" tiuski Matti turhautuneena. Kului muutama viikko ja pake-tillinen laastareita. Ehkä kolmaskin viikko. Isäkin taisi antaa jo hiukan periksi, kun harjoittelu ei tuntunut tuottavan tulosta. Eräänä päivänä isän touhutessa autotallissaan, kuuli hän huudon kadulta: "Katso isä! Minä osaan ajaa polkupyörälläni!" huusi Matti ja viiletti kodin ohi kohti uusia seikkailuja. Onkohan Matilla jo mielessä jokin uusi ta-voite, sellainen, joka tulee mieleen päivittäin?

Carrot

Jänösten perhe asui Pottusen maatilan viereisessä metsässä. Perheen leikkisät lapsoset Pekka ja Elli leikkivät Masa Mäyrän ja Herkko Hiiren kanssa. "Herkkoa on ihan maadoton löytää heinikosta", tuskailivat muut leikkiessään piilosta. Kun eläinlapsoset kaipaavat jännitystä elämäänsä, käyvät he pihistämässä porkkanan tai kaksi Herra Pottusen kasvi-maalta. Se on kiellettyä ja vaarallistakin. "Tällä kertaa aion nostaa ylös sen jätti-porkkanan!" uhosi Pekka voimiensa tunnossa. "Älä luulekaan, sitä tuskin saa irti maasta itse Herra Pottunenkaan", epäilivät muut. Pian lapset olivatkin jo aidan väärällä puolen nykyimässä irti tuota jättiläismäistä porkkanaa. "Huoh, ei se nouse, vaikka vedämme yhdessä tuumin", parkaisivat eläinpolot. Samassa alkoi kuulua askeleita. Lapset pinkaisivat pa-koon, mutta jäivät tarkkailemaan tilannetta aidanraosta. Herra Pottunen oli tullut kor-jaamaan porkkanasatoa. Hän nyppi vaivattomasti porkkanat irti maasta, nakkasi ne sankoon ja lähti pois. Lähtiessään hän kompuroi sen verran, että jättiporkkana tipahti huomaamatta sangosta maahan. "Päivällinen on katettu, kuka uskaltaa tulla syömään kanssani?" totesi aina rauhallinen Masa, muiden vielä täristessä säikähdyksestä.

LIITE B: AINEISTON ARVIOINTIIN KÄYTETYT KRITEERIT

A. Sanojen painotus, ryhmittäminen ja tauotus, kieleen kytkeytyminen, syntaktinen taso eli lauseen sisäinen sanapainojen vaihtelu. (Esim. *“Tule alas sieltä!”*).

1. Toistuvaa sanasta sanaan tai tavusta tavuun lukemista, usein takellellen, monotonisella äänenpainolla. Näin ollen jopa sanatason painotus on säännöllisesti häiriintynyt tai sanat luetaan täysin yksittäin. Lauseen alun/lopun painotukset puuttuvat useimmiten tai ovat enemmänkin huilaustaukoja. Hitaasta lukemisesta johtuen lauseiden välisen tauon pituus ei myöskään tue syntaksia. Tutuimmat sanat ja fraasit voi silti onnistua. Huom. tällainen lukeminen voi olla lähes virheetöntä, mutta hirveän hidasta ja työlästä, ja monotonista.

2. Lauseen alun/lopun painotukset ja lauseiden välinen tauko puuttuvat usein. Myös asiaankuulumatonta painotusta, intonaatiota ja taukoja, jotka eivät osoita kirjoitettujen virkkeiden ja lauseiden alkua ja loppua. Ongelmat painotuksissa erityisesti niillä, jotka lukevat toistuvasti kahden ja kolmen sanan ryppäissä töksähtelevästi, mutta myös sujuvammilla lukijoilla. Todella nopeat pikalukijat voivat olla tätä tasoa, jos sana/tavupainotus hyvin latistunutta tai puuroutunutta.

3. Vaihtelevia lukupyrähdyksiä, hengähdystaukoja kesken lauseiden, jonkin verran töksähtelyä; oikea kielellinen jäsenitys kuultavissa suurimman osan ajasta. Lauseen alussa/lopusa asiallinen painotus ja asiallinen tauko lauseiden välissä. Satunnaisesti kielellinen jäsenitys kuitenkin katoaa. Lukeminen monotonista. Joillakin lukeminen kuulostaa hätäiseltä.

4. Lukeminen on sujuvaa ja ulosanti kielellisesti hyvin jäsentynyttä syntaksin ja merkityksen tasolla. Kuitenkin äänenkäyttö on rajoittunutta verrattuna parhaimpiin lukijoihin. Kyse lienee siitä, että ei ole vain harjaantunut puhumaan/lukemaan kovin selkeästi. 3-ei taas olisi reilu, jos syntaksi on täysin kunnossa. Nelostasoon johtaa merkitystä ilmentävien sanapainotusten vähäisyys lauseen sisällä, tai muutamat selvät puutteet lause- ja virkeyksiköiden huomioimisessa.

5. Yleisesti hyvät painotukset, jotka ovat pääosin lause- ja virkeyksiköissä. Huomioi merkityksen ilmaisun asianmukaisesti selvillä sanapainotuksilla. Tämä kohta on varattu parhaille tekstin merkityksen ilmentäjille. Varo sekoittamasta tunneilmaisuuksiin.

B. Sujuvuus (dekoodaamisen taito, takeltelu, toistelu, itsekorjailu erityisesti hankalien rakenteiden kohdalla).

1. Toistuvia pitkiä taukoja, epäröintiä, varaslähtöjä, äännähdyksiä, toistoja ja/tai useita yrityksiä helppoissakin tekstikohdissa. Tästä syystä eteneminen tekstissä on hyvin hidasta.

2. Toistuvia hankalia kohtia tekstissä, jossa tauot, epäröinnit, toistot ja/tai takertelut esiintyvät häiritsevästi.
3. Useat tauot, epäröinnit, toistot ja/tai takertelut häiritsevät lukemisen sujuvuutta.
4. Satunnaisia taukoja sujuvuudessa johtuen vaikeuksista tiettyjen sanojen ja/tai rakenteiden kohdalla. Ongelmat ratkeavat pian itseoikaisemalla.
5. Yleisesti sujuvaa lukemista, ihan muutamia hankalampia lyhyitä kohtia.

C. Tahti (yleinen etenemisen vauhti)

1. Hidas ja työläs. Usein kuultavissa etenemistä sanatasolla tavuittain/liu'uttamalla.
2. Kohtalaisen hidas tai säännöllisen toistuvasti hidasta. Etenee hitaasti mutta varmasti sana kerrallaan.
3. Rauhallisen sujuva tahti. Voi olla myös epätasaisesti hidasta ja nopeaa lukemista.
4. Reipas lukemisen tahti. Hankalien lausekkeiden osuessa kohdalle lukunopeus kuitenkin hetkellisesti hidastuu.
5. Yleisesti reipas tahti, vaikka mahdollisesti ihan muutamia epäröintejä. Johdonmukaisesti keskustelevaa lukemista.

D. Tunneilmaisuus - Tunnetaso, eläytyminen. Huomaa ero sanojen painotuksen ja tunneilmaisun välillä: Selkeät painotukset (kuten sanapainotus ja lauserajat) eivät vielä vastaa hyvää tunneilmaisua, vaan tunneilmaisua on esimerkiksi repliikin selkeä korostaminen sopivilla äänenpainolla tai äänenkorkeuden vaihteluilla tai repliikin lukeminen tunnetta ilmaisevalla äänensävyllä.

1. Ei tunneilmaisua. Lukee hiljaisella äänellä kuin itselleen tapaillen. Ei kuulosta luonnolliselta siten kuin puhuisi ystävälleen. Tai niin takeltelevaa ettei ilmaisulle jää mahdollisuutta.
2. Kykenee käyttämään tunneilmaisua yksittäisessä repliikissä. Muuten lukeminen ei sisällä tunneilmaisua.
3. Ilmaisullisuutta on jossain määrin sekä repliikeissä että muualla tekstissä. Lukee hiljaisella äänellä. Kuulostaa vain paikoin luonnolliselta puheelta, siten kuin puhuisi ystävälle. Joistain kuulee, että he ymmärtävät tekstin ilmaisullisen tason, mutta heiltä puuttuu vielä ilmaisukeinot. Tällöin ilmaisu voi kuulla hiukan, kun oikein tarkasti kuuntelee, mutta ei olisi reilua antaa heille yhtä hyvin pisteitä kuin niille, jotka saavat ilmaisunsa selkeästi ja rohkeasti ulos.
4. Lukee tekstin pääsääntöisesti ilmeikkäästi. Lukee äänekkäästi ja ilmaisullisesti, mutta paikoitellen ilmaisu häviää ja lukeminen ei enää kuulosta siltä kuin puhuisi ystävälle. Puolet tai kolme neljäsosaa

tekstistä luettiin johdonmukaisella tyyllillä. Verrattuna vitoseen tunneilmaisu on kuitenkin varovaisempaa.

5. Lukee melkein koko tekstin linjassa tarinan kirjoittajan tarkoittaman sävyn mukaisesti. Lukee sovittean ilmaisun ja äänenvoimakkuuden tekstiin sopivasti. Kuulostaa siltä kuin puhuisi ystävälle. Käyttää ääntään ja ilmaisukeinoja todella reippaasti ja rohkeasti.